B&P File No. 9369-161

BERESKIN & PARR **UNITED STATES**

Title: Oil Bodies and Associated Proteins as Affinity Matrices

Inventors: Maurice Moloney, Joseph Boothe and Gijs van Rooijen

TITLE: **Oil Bodies and Associated Proteins as Affinity Matrices**

This application is a continuation in part application of United States application serial no. 09/319,275 filed August 27, 1999 which is a continuation in part of United States application serial no. 08/767,026 filed December 16, 1996, both of which are incorporated herein by reference.

FIELD OF THE INVENTION

This invention relates to the use of oil bodies and their associated proteins as affinity matrices for the separation and purification of target molecules from samples.

BACKGROUND OF THE INVENTION

Within the general field of biotechnology, the ability to effectively separate and purify molecules from complex sources, such as living cells, blood serum, or fermentation broth, is of critical importance. Applications in industry and research laboratories (where, for example, purified or partly purified proteins are used) are numerous and well documented in prior literature. See, for example, R. Meadon and G. Walsh in *Biotechnological Advances* 1994, **12**: pp 635-646.

The majority of currently employed techniques for the separation of molecules capitalizes on the innate physical and chemical properties of the molecule of interest. Affinity-based purification technologies are unique in that they exploit the highly specific biological recognition between two molecular species which form an affinity pair. Binding of the two entities of the affinity pair occurs in almost all instances as a result of relatively weak chemical interactions, known as non-covalent bonds. Some art-recognized and commonly used affinity pairs include antibodies and their binding antigenic substances, nucleic acid binding proteins and nucleic acids, lipid binding proteins and lipids, lectins and carbohydrates, streptavidin/biotin complexes, protein A/immunoglobulin G complexes, and receptors and their binding molecules.

In general, affinity-based purification processes require that one member of the affinity pair is immobilized on a solid substrate or matrix that is insoluble in the fluid in which the other member of the pair

resides. The molecular species of the affinity pair bound to the matrix is generally referred to as the ligand, while the liquid soluble member is generally referred to as the target member. However, it is important to note that these definitions do not impose any restrictions in a strict chemical sense. The vast majority of current ligand immobilization techniques rely on physical or chemical approaches. Physical ligand immobilization involves adsorption or entrapment of the ligand to a suitable support, while the chemical mode of immobilization is characterized by the formation of strong crosslinks or covalent attachments between the ligand and the matrix. It is a requirement that immobilization is accomplished in such a fashion that the capacity of the members of the affinity pair to recognize each other is not adversely affected by the immobilization procedure.

It is a disadvantage of the currently available physical and chemical techniques for immobilizing ligands that production processes are frequently time consuming and expensive. This is mainly due to the fact that immobilization techniques require the separate production of matrix material and ligands, which in a subsequent step must be coupled. An alternative mode of immobilizing proteins is described in U.S. Patent No. 5,474,925 which documents a biological production system for the immobilization of enzymes in the fibre of cotton plants. This patent discloses what is believed to be the first biologically produced enzyme immobilization system and allows a one step production of matrix and ligand.

Subsequent to immobilization of the ligand on the matrix, a variety of affinity based purification techniques may be employed to accomplish selective binding between the affinity immobilized ligand and the target member. Affinity based purification techniques known in the prior art include perfusion affinity chromatography, affinity repulsion chromatography, hyperdiffusion affinity chromatography, affinity precipitation, membrane affinity partitioning, affinity cross-flow ultrafiltration and affinity precipitation. In the most widely used affinity based purification technique, affinity chromatography, a matrix containing a ligand is coated to, or packed on, the inside of a chromatographic column. A complex mixture containing the target member is then applied to the

chromatographic column. Ideally, only the target molecules that specifically recognize the ligand bind in a non-covalent fashion to the chromatographic column, while all other molecular species present in the sample pass through the column.

In affinity partitioning, two solutions of substantially different densities are employed. The complex solution containing the target member is mixed with a solution of a different density containing the affinity ligand. Subsequent to mixing, the solutions are left to settle in order to permit the formation of two separate phases. Molecules tend to partition differentially between phases depending on their size, charge and specific interactions with the phase-forming solutions. Ligand-bound target protein selectively partitions to the phase containing the affinity ligand. For example, Coughlin and Baclaski in *Biotechnology Progress*, 1990 6: 307-309 reported the use of the biotin containing organic solution isooctane to transfer avidin from an aqueous solution to the isooctane solution. However, so far applications of affinity partitioning have been limited mainly due to the current lack of availability of suitable affinity matrix substances which can be employed in specific partitioning in two phase systems.

An important factor for the commercial development of biotechnology is the purification of bioproducts, which typically accounts for 50% or more of the total costs (Labrou, N. and Clonis, Y. D. in the *Journal of Biotechnology* 36: 95-119 (1994)). Many protein purification steps rely on column type separation procedures. In particular, large scale high-separation techniques such as column chromatography or batch-type based

processes have been reviewed by Labrou, N. and Clonis, Y. D. in the *Journal of Biotechnology* 36: 95-119 (1994).

There is a need in the art to develop novel and economical methods for separating and purifying biological products from complex mixtures. The present inventors have found that subcellular oil storage structures, known as oil bodies, and their associated proteins are useful in this regard.

SUMMARY OF THE INVENTION

The present invention relates to a novel versatile biological system for the production of affinity matrices. The present inventors have found that oil bodies and their associated proteins can be used as affinity matrices for the separation of a wide variety of target molecules such as proteins, carbohydrates, lipids, organic molecules, nucleic acids, metals, cells and cell fractions from a sample.

In accordance with the invention, there is provided a method for the separation of a target molecule from a sample comprising: 1) contacting (i) oil bodies that can associate, either directly or indirectly, with the target molecule with (ii) a sample containing the target molecule; and 2) separating the oil bodies associated with the target molecule from the sample. The oil bodies and the sample containing the target molecule are brought into contact in a manner sufficient to allow the oil bodies to associate with the target. Preferably, oil bodies are mixed with the target. If desired, the target molecule may be further separated from the oil bodies.

In one aspect, the target molecule has affinity for, or binds directly to, the oil bodies or oil body protein. Examples of such targets include antibodies or other proteins that bind to oil bodies.

In another aspect, a ligand molecule may be used to associate the target molecule with the oil bodies.

In one embodiment, the ligand has natural affinity for the oil bodies or oil body protein. In a specific embodiment, the ligand is an antibody that binds the oil body protein. Such an antibody can be used to separate targets having natural affinity for the ligand antibody such as anti-IgG antibodies or protein A. A bivalent antibody may also be prepared having binding specificity for both the oil body protein and the target. The

antibody against the oil body protein may also be fused to a second ligand having affinity for the target. The antibody against the oil body protein may also be covalently attached to the target molecule. In one embodiment, the target molecule can be a protein that is produced as a recombinant fusion protein with the ligand antibody.

In another embodiment, the ligand is covalently attached to the oil bodies or oil body protein. In one embodiment, the ligand is a protein that is chemically conjugated or produced as a fusion protein with the oil body protein (as described in WO 96/21029). In the latter case, the fusion protein is targeted to and expressed on the oil bodies. In one example, the ligand fused to the oil body protein may be hirudin and can be used to purify thrombin. In another example, the ligand fused to the oil body protein may be metallothionein and can be used to separate cadmium from a sample. In a further example, the ligand fused to the oil body protein may be protein A and can be used to separate immunoglobulins. In yet another example, the ligand fused to the oil body protein may be cellulose binding protein and can be used to separate cellulose from a sample.

In another embodiment, the ligand may be covalently attached to the oil bodies. For example, the ligand may be a small organic molecule such as biotin. Biotinylated oil bodies can be used to separate avidin from a sample.

The present invention also includes modified oil bodies for use as an affinity matrix. Accordingly, the present invention includes a composition comprising oil bodies associated with a molecule, such as a ligand molecule or a target molecule. In one embodiment, the composition comprises oil bodies covalently attached to a ligand molecule, such as biotin.

The present invention also includes an affinity matrix for use in separating a target molecule from a sample, comprising oil bodies that can associate with the target molecule. The affinity matrix may additionally include a ligand molecule associated with the oil bodies, wherein the ligand molecule is capable of associating with the target molecule.

Other objects, features and advantages of the present invention will become apparent from the following detailed description and

attached drawings. It should be understood, however, that the detailed description and associated examples are given by way of illustration only, and various changes and modifications thereto falling within the scope of the invention will become apparent to those skilled in the art. In addition, reference is made herein to various publications, patents and patent applications which are hereby incorporated by reference in their entirety.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. The nucleotide and deduced amino acid sequence of the 18 KDa oleosin from *Arabidopsis thaliana* as shown in SEQ.ID.NO:1 and SEQ.ID.NO:2.

Figure 2. Sequence of an *Arabidopsis* oleosin-hirudin fusion. Indicated are a portion of the oleosin genomic sequence (from base 1-1620 as reported in van Rooijen *et al* 1992, *Plant Mol. Biol*. 18: 1177-1179), a spacer sequence (base 1621-1635, underlined) and the synthetic DNA sequence encoding the mature hirudin variant-2 isoform (base 1636-1833, italicized) This gene fusion is regulated by the 5′ upstream region of the *Arabidopsis* oleosin (bases 1-861) and the nopaline synthase termination sequence (base 1855-2109). The sequence is also shown in SEQ.ID.NO:3 and SEQ.ID.NO:4.

Figure 3. Outline of the steps employed in the construction of pCGOBHIRT, containing the entire oleosin-hirudin construct.

Figure 4. Schematic diagram illustrating the configuration of the oleosin-hirudin fusion protein on the oil body and the binding of thrombin.

Figure 5. NaCl elution profiles of thrombin from wild type and 4A4 oil body matrices transformed with a construct expressing an oleosin-hirudin fusion.

Figure 6. Purification of a horseradish peroxidase conjugated anti-IgG antibody using an anti-oleosin antibody as a ligand. Schematic diagram illustrating the configuration of the oleosin / anti-oleosin / anti-IgG sandwich complex bound to an oil body.

Figure 7. Illustrates specific binding of anti-IgG antibodies to wild type oil bodies complexed with primary anti-oleosin antibodies as a ligand (left) and binding of anti-IgG antibodies to oil bodies which were not

complexed with primary antibodies prior to binding with the secondary antibodies (right).

Figure 8. Sequence of an oleosin metallothionein fusion. Indicated are the coding sequence of a *B. napus* oleosin cDNA (bases 1092-1652, van Rooijen, 1993, *Ph.D. Thesis*, University of Calgary), a spacer sequence (bases 1653-1670, underlined) and the human metallothionein gene *mt-II* (bases 1671-1876, Varshney and Gedamu, 1984, *Gene*, **31**: 135-145)). The gene fusion is regulated by an *Arabidopsis* oleosin promoter (bases 1-1072) and ubiquitin termination sequence (bases 1870-2361, ubi3'; Kawalleck *et al.*, 1993, *Plant Mol. Biol.* **21**: 673-684). The sequence is also shown in SEQ.ID.NO:6 and SEQ.ID.NO:7.

Figure 9. Outline of the steps employed in the construction of pBIOOM3' containing the entire oleosin-metallothionein construct.

Figure 10. Schematic diagram illustrating the configuration of the oleosin-metallothionein fusion protein on the oil body and binding of cadmium ions.

Figure 11. Illustrates the binding (A) and elution (B) of cadmium to an oil body matrix from wildtype *B. carinata* seeds and *B. carinata* seeds transformed with a construct expressing oleosin metallothionein gene fusion. Shown is the percentage cadmium bound to the oil body fraction of an oil body fraction harvested from transgenic and untransformed seeds. Bars represent average values of 5 replicate experiments (binding) and 3 replicates (elution).

Figure 12. Illustrates the binding of protein A expressing *S. aureus* cells to oil bodies treated with varying amounts of anti-oleosin IgGs. Bars represent OD_{600} readings obtained following the procedures as described in Example 5 and using varying amounts of IgGs (0 µl, 3, µl, 30 µl, 100 µl of added IgG).

Figure 13. Oligonucleotide primers used to amplify the sequence of the *S. aureus* protein A (The sequence is also shown in SEQ.ID.NO:8; The protein sequence is also shown in SEQ.ID.NO:9). Primer BK266, 5'C *TCC ATG* GAT CAA CGC AAT GGT TTA TC 3' (SEQ.ID.NO:10), a *NcoI* site (italicized) and a sequence identical to a portion of the protein A gene as contained within vector pRIT2T (Pharmacia) (underlined) are

indicated. Primer BK267, 5' GC *AAG CTT* **CTA** ATT TGT TAT CTG CAG GTC 3' (SEQ.ID.NO:11), a *HindIII* site (italicized), a stop codon (bold) and a sequence complementary to a portion of the protein A gene as contained within pRIT2T (Pharmacia) (underlined) are indicated. The PCR product was digested with *NcoI* and *HindIII* and ligated into pCGNOBPGUSA (Van Rooijen and Moloney, 1995, *Plant Physiol.* **109**: 1353-1361) from which the *NcoI*-GUS-*HindIII* fragment had been removed.

Figure 14. Sequence of an *Arabidopsis* oleosin-protein A fusion (The sequence is also shown in SEQ.ID.NO:12 and the protein sequence is also shown in SEQ.ID.NO:13 and 14). Indicated are a portion of the oleosin genomic sequence (from base 1 - 1626, as reported in van Rooijen *et al.*, 1992 *Plant Mol. Biol.* **18**: 1177-1179), a spacer sequence encoding a thrombin cleavage site (base 1627 - 1647, underlined) and the DNA sequence encoding protein A (base 1648 - 2437, italicized). Expression is regulated by the *Arabidopsis* 5' upstream region of the *Arabidopsis* oleosin (base 1 - 867) and the nopaline synthase terminator region (base 2437 - 2700).

Figure 15. Schematic diagram illustrating the configuration of the oleosin-protein A fusion protein on the oil body and binding of the immunoglobulin.

~~**Figure 16.** A western blot illustrating the binding of HRP-conjugated mouse anti-rabbit antibodies to oil body protein extracts obtained from transgenic *B. napus* lines expressing oleosin-protein A fusion proteins. Shown on a Western blot probed with an HRP- conjugated antibody are oil body protein extracts from transgenic lines, opa 30 (lane 3), opa 31 (lane 4), opa 34 (lane 5), opa 36 (lane 6), opa 47 (lane 7), opa 93 (lane 8), all expressing an oleosin-protein A fusion protein and a control untransformed *B. napus* line (lane 9), as well as lysates of *E. coli* DH5α transformed with pRIT2T expressing protein A (lane 2) and untransformed *E. coli* DH5α (lane 1).~~

Figure 17. illustrates binding and elution of IgGs to oil bodies isolated from wildtype *B. napus* (bn wt) and a transgenic *B. napus* line,

Figures 18A-C. Nucleotide sequence of the phaseolin promoter-PRS-OBScFv-Prochymosin -phaseolin terminator sequence (SEQ ID No. 15). The phaseolin promoter corresponds to nucleotide 6-1554. The DNA sequence encoding the PRS-OBScFv-Prochymosin gene fusion corresponds to nt 1554-3467. The phaseolin terminator corresponds to nucleotide sequence 3474-4694. The deduced amino acid sequence of the PRS-OBScFv-Prochymosin fusion is also indicated. Met1 to Ala25 corresponds to the PRS signal sequence, Glu28 to Thre 142 corresponds to the Variable heavy antibody chain, Gly 143 to Ser156 corresponds to the flexible linker peptide separating Vh and Vl, Asp157-Leu271 corresponds to the Variable light chain and Ala273 to Ile 638 corresponds to the prochymosin peptide.

Figure 19. Western Blot of seed and oil body extracts of *Arabidopsis* plants transformed with pSBS2168. A total of 12.5 µl of sample was loaded in each lane. Lane 1; Bovine chymosin, Lane 2; total extract, Lane 3; Sup1, Lane 4; OB0, Lane 5 ;OBHighS. See the text for a description of the samples. This Western blot was treated with polyclonal antibodies raised against bovine chymosin followed by an alkaline phosphatase linked secondary antibody and NBT/BCIP color reaction. Indicated are the OBScFvProchymosin protein fusion (OBScFvProchymosin), a band which could correspond to an aberrantly folded OBScFvProchymosin protein fusion, and as a result has a slightly slower mobility on the polyacrylamide gel and no detectable affinity to *Arabidopsis* oil bodies. Also indicated are the bands that correspond to the mature processed form of chymosin (chymosin) and proteolytic breakdown products of chymosin (partial chymosin).

DETAILED DESCRIPTION OF THE INVENTION

As hereinbefore mentioned, the present invention relates to a novel biological affinity matrix system that employs oil bodies and their associated proteins. The affinity matrix is suitable for the highly-efficient separation of specific targets, including proteins, carbohydrates, lipids, nucleic acids, cells and subcellular organelles, metals and ions, from a sample.

The present invention provides a method for the separation of a target molecule from a sample comprising: 1) contacting (i) oil bodies that can associate either directly or indirectly with the target molecule with (ii) a sample containing the target molecule; and 2) separating the oil bodies associated with the target molecule from the sample. The oil bodies and the sample containing the target molecule are brought into contact in a manner sufficient to allow the oil bodies to associate with the target. Preferably, the oil bodies are mixed with the target. Indirect association of the oil bodies with the target can be effected using a ligand molecule that can associate with both the oil bodies and the target molecule. The ligand therefore serves to bridge or join the oil bodies with the target molecule. If desired, the target molecule may be further separated from the oil bodies and the ligand, if present.

Each of the components of the affinity matrix are discussed in turn below.

Targets

The term "target" as used herein denotes a desired molecule that one wants to purify, isolate or separate from a sample such as a biological mixture. This technology is amenable for use with virtually any target for which a ligand can be obtained or any target that can directly associate with or bind to an oil body or oil body protein. Possible ligand/target pairs include but are not limited to: protein subunit/subunit associations, antibodies/antigens, receptor protein/signal molecules, nucleic acid binding proteins/nucleic acids; lectins/carbohydrates; lipid binding proteins/lipids; ion binding proteins/ions; and ligands to surface epitopes/cells or subcellular organelles. The target may be obtained from any natural source or may be synthesized chemically. If the target is a macromolecule such as a protein or nucleic acid it may also be produced in recombinant form using any suitable expression system such as bacteria, yeast, plant, insect, mammalian, etc.

Ligands

The term "ligand" used herein denotes a molecule that is capable of associating with both the target molecule and the oil bodies or oil body protein (discussed below). The term "associating with" as used he

includes both covalent and non-covalent binding of the ligand to the oil bodies or the target molecule. For example, the ligand molecule may be covalently attached to the oil bodies (or oil body protein) and non-covalently associate with the target (and vice-versa), or the ligand may non-covalently associate with both the oil bodies and the target molecule. The ligand may be any molecule that can bridge the oil bodies or oil body protein and the target molecule and can include a protein, nucleic acid, carbohydrate or small organic molecule. The ligand may be comprised of two molecules, a first molecule that associates with the oil bodies and a second molecule that associates with the target, wherein the first molecule and the second molecule associate with each other.

The affinity ligand proteins used for this methodology may be derived from naturally-occurring, known ligand pairs such as those listed above. Alternatively, the ligand may be obtained by screening proteins extracted from cells or organisms, synthesized chemically or produced in libraries comprised of combinatorial peptide sequences, antibodies, or expressed DNA sequences.

In one embodiment, the ligand has natural affinity for the oil bodies or the oil body protein. For example, the ligand may be a protein such as an antibody, that has affinity for the oil body protein. The ligand may also be a molecule other than a protein which has natural affinity for the oil body or oil body protein. Such ligands, capable of binding to the oil bodies or oil body protein, may be associated either directly or indirectly with the target molecule. In a particular embodiment, the ligand is covalently attached to the target molecule by chemical or recombinant means. For example, the ligand may be an antibody that is prepared as a recombinant fusion protein with the target. The ligand may also be associated with a second molecule that can bind the target molecule. For example, the ligand molecule may be an antibody conjugated to avidin and can be used to purify biotin from a sample.

In another embodiment, the ligand is covalently linked to the oil bodies or oil body protein by chemical or recombinant means. Chemical means for preparing fusions or conjugates are known in the art and can be used to prepare a ligand-oil body protein fusion. The method used to

conjugate the ligand and oil body must be capable of joining the ligand with the oil body protein without interfering with the ability of the ligand to bind to the target molecule. In one example, the ligand may be a small organic molecule such as biotin that is covalently attached to the oil bodies. Biotinylated oil bodies can be used to separate avidin from a sample. The present invention also includes modified oil bodies such as biotinylated oil bodies for use as an affinity matrix. Accordingly, the present invention includes a composition comprising oil bodies attached to a molecule, such as a ligand or a target molecule.

In a preferred embodiment, the ligand is a protein and can be conjugated to the oil body protein using techniques well known in the art. There are several hundred crosslinkers available that can conjugate two proteins. (See for example "Chemistry of Protein Conjugation and Crosslinking". 1991, Shans Wong, CRC Press, Ann Arbor). The crosslinker is generally chosen based on the reactive functional groups available or inserted on the ligand. In addition, if there are no reactive groups a photoactivatible crosslinker can be used. In certain instances, it may be desirable to include a spacer between the ligand and the oil-body protein. Crosslinking agents known to the art include the homobifunctional agents: glutaraldehyde, dimethyladipimidate and Bis(diazobenzidine) and the heterobifunctional agents: *m*-Maleimidobenzoyl-*N*-Hydroxysuccinimide and Sulfo-*m*-Maleimidobenzoyl-*N*-Hydroxysuccinimide.

A ligand protein-oil body protein fusion may also be prepared using recombinant DNA techniques. In such a case a nucleic acid sequence encoding the ligand is fused to a nucleic acid sequence encoding the oil body protein, resulting in a chimeric nucleic acid molecule that expresses a ligand-oil body protein fusion protein (discussed in greater detail below). In order to prepare a recombinant fusion protein, the sequence of the nucleic acid encoding the ligand must be known or be obtainable. By obtainable it is meant that a nucleic acid sequence sufficient to encode the protein ligand may be deduced from the known amino acid sequence. It is not necessary that the entire gene sequence of the ligand be used provided that a

Therefore, the ligand can include the complete sequence of, or the binding domain from, the specific ligand protein in question.

If the nucleic acid sequence of the desired ligand is known, the gene may be synthesized chemically using an oligonucleotide synthesizer. Alternatively, the clone carrying the ligand gene may be obtained from either cDNA or genomic libraries containing the gene by probing with a labelled complementary nucleic acid sequence. The gene may also be specifically amplified from the library using gene-specific oligonucleotide primers and the PCR. If the nucleic acid sequence of the desired ligand is not known, then a partial amino acid sequence may be obtained through N-terminal sequencing of the protein (Matsudaira 1987; *J. Biol. Chem.* 262: 10035-10038). Labelled probes may be synthesized based upon the DNA sequences deduced from this amino acid sequence and used to screen cDNA or genomic libraries as described above. The clone carrying the gene may also be identified from a cDNA expression library by probing either with antibodies raised against the protein ligand, or with the target protein.

Ligands may also be uncovered by probing mixtures of proteins with the target. The target can be immobilized on a support matrix and used to screen proteins extracted from cells and tissues or synthesized chemically. Following binding between the ligand protein and the immobilized target, the matrix is separated from the solution and washed. The protein ligand is subsequently eluted from the matrix and the sequence determined as described above. Alternatively, recombinant protein libraries produced by phage display, such as those comprised of combinatorial peptide sequences (Smith, 1985; *Science* 228: 1315-1317) or antibody repertoires (Griffiths *et al.*, 1994, *EMBO J.* 13: 3245-3260, Nissim *et al.*, 1994, *EMBO J.* 13: 692-698) can be screened with the immobilized target. In this

construct are transformed with constructs from a cDNA library where the sequences have been fused to that of a transcriptional activator. Binding between ligands derived from the cDNA library with the target protein allows transcription of a reporter gene to occur. Clones expressing the ligand are then recovered.

To specifically uncover a ligand to oil bodies, a complete or partial oleosin protein may be used as target in any of the above methods. Alternatively, it may be possible to employ intact oil bodies for screening protein extracts, synthetic peptides or phage display libraries. In this case, the oil body would serve both as target and immobilization matrix. Using this approach, a wider variety of ligands may be uncovered; that exhibit affinity not only to oleosins, but to other epitopes present on oil bodies.

__Oil bodies and Oil Body Proteins__

Oil bodies are small, spherical, subcellular organelles encapsulating stored triacylglycerides, an energy reserve used by many plants. Although they are found in most plants and in different tissues, they are particularly abundant in the seeds of oilseeds where they range in size from under one micron to a few microns in diameter. Oil bodies are comprised of the triacylglycerides surrounded by a half-unit membrane of phospholipids and embedded with a unique type of protein known as an oil body protein. The term "oil body" or "oil bodies" as used herein includes any or all of the triacylglyceride, phospholipid or protein components present in the complete structure. The term "oil body protein" as used herein means a protein that is naturally present in an oil body. In plants, the predominant oil body proteins are termed "oleosins". Oleosins have been cloned and sequenced from many plant sources including corn, rapeseed, carrot and cotton. The oleosin protein appears to be comprised of three domains; the two ends of the protein, N- and C-termini, are largely

Oil bodies from individual species exhibit a roughly uniform size and density which is dependent in part upon the precise protein/phospholipid/triacylglyceride composition. As a result, they may be simply and rapidly separated from liquids of different densities in which they are suspended. For example, in aqueous media where the density is greater than that of the oil bodies, they will float under the influence of gravity or applied centrifugal force. In 95% ethanol where the density is less than that of the oil bodies, they will sediment under the same conditions. Oil bodies may also be separated from liquids and other solids present in solutions or suspensions by methods that fractionate on the basis of size. For example, the oil bodies from *B. napus* are minimal, approximately 0.5μm in diameter, and thus may be separated from smaller components using a membrane filter with a pore size less than this diameter.

The oil bodies of the subject invention are preferably obtained from a seed plant and more preferably from the group of plant species comprising: thale cress (*Arabidopsis thaliana*), rapeseed (*Brassica* spp.), soybean (*Glycine max*), sunflower (*Helianthus annuus*), oil palm (*Elaeis guineeis*), cottonseed (*Gossypium* spp.), groundnut (*Arachis hypogaca*), coconut (*Cocus nucifera*), castor (*Ricinus communis*), safflower (*Carthamus tinctorius*), mustard (*Brassica* spp. and *Sinapis alba*), coriander (*Coriandrum sativum*) linseed/flax (*Linum usitatissimum*), and maize (*Zea mays*). Plants are grown and allowed to set seed using agricultural cultivation practises well known to a person skilled in the art. After harvesting the seed and removal of foreign material such as stones or seed hulls, for by example sieving, seeds are preferably dried and subsequently processed by mechanical pressing, grinding or crushing. The oil body fraction may be obtained from the crushed seed fraction by capitalization on separation techniques which exploit differences in density between the oil body fraction and the aqueous fraction, such as centrifugation, or using size exclusion-based separation techniques, such as membrane filtration, or a combination of both of these. Typically, seeds are thoroughly ground in five volumes of a cold aqueous buffer. A wide variety of buffer compositions may be employed, provided that they do not contain high concentrations of strong organic solvents such as acetone or diethyl ether, as these solvents may disrupt the oil bodi

solution density of the grinding buffer may be increased with the addition of 0.4-0.6 M sucrose, in order to facilitate washing as described below. The grinding buffer will also typically contain 0.5 M NaCl to help remove soluble proteins that are not integrally bound to the oil body surface.

Following grinding, the homogenate is centrifuged resulting in a pellet of particulate and insoluble matter, an aqueous phase containing soluble components of the seed, and a surface layer comprised of oil bodies with their associated proteins. The oil body layer is skimmed from the surface and thoroughly resuspended in one volume of fresh grinding buffer. It is important that aggregates of oil bodies are dissociated as thoroughly as possible in order to ensure efficient removal of contaminants in the subsequent washing steps. The resuspended oil body preparation is layered under a floatation solution of lower density (e.g. water, aqueous buffer) and centrifuged, again, separating oil body and aqueous phases. This washing procedure is typically repeated at least three times, after which the oil bodies are deemed to be sufficiently free of contaminating soluble proteins as determined by gel electrophoresis. It is not necessary to remove all of the aqueous phase and to the final preparation water or 50 mM Tris-HCl pH 7.5 may be added and if so desired the pH may be lowered to pH 2 or raised to pH 10. Protocols for isolating oil bodies from oil seeds are available in Murphy, D. J. and Cummins I., 1989, Phytochemistry, **28**: 2063-2069; and in: Jacks, T. J. et al., 1990, JAOCS, **67**: 353-361. A preferred protocol is detailed in example 1 of the present specification.

Oil bodies other than those derived from plants may also be used in the present invention. A system functionally equivalent to plant oil bodies and oleosins has been described in bacteria (Pieper-Fürst et al., 1994, *J. Bacteriol.* **176**: 4328), algae (Rossler, P.G., 1988, *J. Physiol. (London)*, **24**: 394-400) and fungi (Ting, J. T. et al., 1997, *J. Biol Chem.* **272**: 3699-3706). Oil bodies from these organisms, as well as those that may be discovered in other living cells by a person skilled in the art, may also be employed according to the subject invention.

<u>Affinity Matrices</u>

As hereinbefore mentioned, the present invention provides a

sample. In one embodiment, the affinity matrix comprises oil bodies that can bind a target molecule in a sample. In such an embodiment, the target molecule may be an antibody that can bind an oil body protein. In another embodiment, the affinity matrix comprises oil bodies or oil body proteins and a ligand that is associated with the oil bodies or oil body proteins and has affinity for a target molecule. In such an embodiment, the ligand may be non-covalently or covalently attached to the oil bodies or oil body protein (as described above). In another embodiment, the affinity matrix comprises oil bodies or oil body proteins and a ligand that is covalently or non-covalently attached to the target.

It is an advantage of the present invention that target substances can be purified or removed from samples through non-covalent association with oil bodies followed by oil body separation. A number of different oil body-ligand configurations are possible. Targets with inherent affinity for a specific ligand proteins such as hirudin to thrombin or heavy metals to metallothionein, may be purified or separated with oil bodies containing that ligand fused to an oleosin. Alternatively, a protein target may also be purified or separated with an oil body affinity matrix by fusing the target to an oil body-specific ligand (such as an antibody that binds the oil bodies or oil body proteins) or to a ligand complimentary to that fused to an oleosin. If desired, a protease recognition site or chemical cleavage site may be engineered between the ligand and the target protein to enable proteolytic removal of the ligand from the target protein in the course of purification. A multivalent ligand may also be constructed, such as a bivalent single-chain antibody, in which one domain of the ligand has an affinity for an oil body and the other domain(s) exhibits affinity for the target. In this case, neither the oil body nor the target molecule need to be covalently fused to a ligand. Also, concatamers of ligands may be used to increase the affinity of a matrix for a target, or the sequence of a ligand may be mutated to modulate the affinity for a target when such conditions are desirable. Further, mixtures of different ligands may be fused to recover/remove different types of targets simultaneously. Fusions

matrix. Binding to the affinity matrix may also be achieved by forming bridges between ligand or ligand and target sequences, such as Zn^{++} ions bridging between polyhistidine sequences.

There are several advantages associated with the use of oil body affinity matrices that make them attractive as purification tools. The flexibility in design that is possible through the different configurations described above, enables a matrix to be constructed to best meet the requirements for a specific target. Also, production of the matrix as part of a natural biological process in seeds is extremely cost-effective, since purification and immobilization of the ligand are not necessary. In the case of oleosin-ligand fusions, the ligand is immobilized on the oil body as a result of oleosin targeting within the cell, while oil body-specific ligands will naturally associate with the matrix while present in complex mixtures. Natural immobilization of the ligand on the matrix may also be advantageous in that it eliminates the requirement for chemical cross-linking that may compromise the affinity of the ligand for the target. Finally, oil body affinity matrices offer a unique and attractive purification option particularly for large scale operations. The ability to separate the matrix through floatation as a loose suspension enables it to be employed with crude material containing what might otherwise be prohibitive amounts of particulate contaminants. The presence of these contaminants will often foul and block conventional solid matrices applied in columns or batch suspensions limiting their use at early stages in the purification process.

As mentioned previously, in one embodiment of the invention, ligand protein sequences are genetically fused to the oil body protein. In order to prepare such genetic fusions, a chimeric nucleic acid sequence is prepared that encodes an oil body protein-ligand fusion protein and comprises (a) a nucleic acid sequence encoding a sufficient portion of an oil body protein to provide targeting of the fusion protein to the oil bodies and (b) a nucleic acid sequence encoding a sufficient portion of the ligand protein to provide binding of the target. The inventors have determined that, in general, the N-terminus and the hydrophobic core of an oil body protein are sufficient to provide targeting of the fusion protein to the oil

bodies. In particular, for oleosins derived from the plant *Arabidopsis thaliana* amino acids 2 through 123 (as shown in SEQ.ID.NO:1) are sufficient in this regard.

The ligand may be fused to either the N- and/or C-terminal end of the oleosin. It may also be possible to construct an internal fusion between the ligand and oleosin or to fuse the ligand between two oleosin proteins. The chimeric DNA sequence encoding an oil body protein fused to a ligand may be transfected into a suitable vector and used to transform a plant. Two types of vectors are routinely employed. The first type of vector is used for the genetic-engineering and assembly of constructs and typically consists of a backbone such as found in the pUC family of vectors, enabling replication in easily-manipulated and maintained gram negative bacteria such as *E. coli*. The second type of vector typified by the Ti and Ri plasmids, specify DNA transfer functions and are used when it is desired that the constructs be introduced into the plant and stably integrated into its genome via *Agrobacterium* mediated transformation.

A typical construct consists, in the 5′ to 3′ direction, of a regulatory region complete with a promoter capable of directing expression in plants (preferably seed-specific expression), a protein coding region, and a sequence containing a transcriptional termination signal functional in plants. The sequences comprising the construct may be either natural or synthetic or any combination thereof.

(1) ~~Both non-seed specific promoters, such as the 35-S CaMV~~ promoter (Rothstein *et al.*, 1987, *Gene* **53**: 153-161) and seed-specific promoters such as the phaseolin promoter (Sengupta-Gopalan *et al.*, 1985; PNAS USA **82**: 3320-3324) or the Arabidopsis 18 kDa oleosin (Van Rooijen *et al.*, 1992; Plant Mol. Biol. **18**: 1177-1179) promoters may be used. In addition to the promoter, the regulatory region contains a ribosome binding site enabling translation of the transcripts in plants and may also contain one or more enhancer sequences, such as the AMV leader (Jobling and Gehrke ~~1987; Nature 325: 622-625), to increase the expression of product.~~

The coding region of the construct will typically be comprised of sequences encoding a ligand fused in frame to an oleosin and ending with a translational termination codon. The sequence for the oleosin may be

Ad(2)

comprised of any DNA sequence, or part thereof, natural or synthetic, sufficient to encode a protein that can be correctly targeted to, and stably expressed on, an oil body. A detailed description of the characteristics of such a sequence has been reported previously in Moloney, 1993; *PCT Patent Appl.* WO 93/21320 which is hereby incorporated by reference. The sequence may also include introns. The ligand-encoding region may in turn be comprised of any individual, or combination of, ligand sequences identified as described above. If desired, a protease or chemical recognition site may be engineered between the ligand and the target protein to enable proteolytic removal of the ligand from the target protein in the course of purification.

The region containing the transcriptional termination signal may comprise any such sequence functional in plants such as the nopaline synthase termination sequence and additionally may include enhancer sequences to increase the expression of product.

The various components of the construct are ligated together using conventional methods, typically into a pUC-based vector. This construct may then be introduced into an *Agrobacterium* vector and subsequently into host plants, using one of the transformation procedures outlined below.

A variety of techniques are available for the introduction of DNA into host cells. For example, the chimeric DNA constructs may be introduced into host cells obtained from dicotyledonous plants, such as tobacco, and oleaginous species, such as *B. napus* using standard *Agrobacterium* vectors; by a transformation protocol such as that described by Moloney *et al.*, 1989, (*Plant Cell Rep.*, 8: 238-242) or Hinchee *et al.*, 1988, (*Bio/Technol.*, 6: 915-922); or other techniques known to those skilled in the art. For example, the use of T-DNA for transformation of plant cells has received extensive study and is amply described in EPA Serial No. 120,516; Hoekema *et al.*, 1985, (Chapter V, In: *The Binary Plant Vector System* Offset-drukkerij Kanters B.V., Alblasserdam); Knauf, *et al.*, 1983, (*Genetic Analysis of Host Range Expression by Agrobacterium*, p. 245, In Molecular Genetics of the Bacteria-Plant Interaction, Puhler, A. ed., Springer-Verlag, NY); and An *et al.*, 1985, (*EMBO J.*, 4: 277-284). Conveniently, explants may be cultivated with

A. tumefaciens or *A. rhizogenes* to allow for transfer of the transcription construct to the plant cells. Following transformation using *Agrobacterium* the plant cells are dispersed in an appropriate medium for selection, subsequently callus, shoots and eventually plantlets are recovered. The *Agrobacterium* host will harbour a plasmid comprising the vir genes necessary for transfer of the T-DNA to the plant cells. For injection and electroporation, (see below) disarmed Ti-plasmids (lacking the tumour genes, particularly the T-DNA region) may be introduced into the plant cell.

The use of non-*Agrobacterium* techniques permits the use of the constructs described herein to obtain transformation and expression in a wide variety of monocotyledonous and dicotyledonous plants and other organisms. These techniques are especially useful for species that are intractable in an *Agrobacterium* transformation system. Other techniques for gene transfer include biolistics (Sanford, 1988, *Trends in Biotech.*, 6: 299-302), electroporation (Fromm *et al.*, 1985, *Proc. Natl. Acad. Sci. USA*, 82: 5824-5828; Riggs and Bates, 1986, *Proc. Natl. Acad. Sci. USA* 83: 5602-5606) or PEG-mediated DNA uptake (Potrykus *et al.*, 1985, *Mol. Gen. Genet.*, 199: 169-177).

In a specific application, such as to *B. napus*, the host cells targeted to receive recombinant DNA constructs typically will be derived from cotyledonary petioles as described by Moloney *et al.*, (1989, *Plant Cell Rep.*, 8: 238-242). Other examples using commercial oil seeds include cotyledon transformation in soybean explants (Hinchee *et al.*, 1988, *Bio/Technology*, 6: 915-922) and stem transformation of cotton (Umbeck *et al.*, 1981, *Bio/Technology*, 5: 263-266).

Following transformation, the cells, for example as leaf discs, are grown in selective medium. Once shoots begin to emerge, they are excised and placed onto rooting medium. After sufficient roots have formed, the plants are transferred to soil. Putative transformed plants are then tested for presence of a marker. Southern blotting is performed on genomic DNA using an appropriate probe, for example an *A. thaliana* oleosin gene, to show that integration of the desired sequences into the host cell genome has occurred.

The expression cassette will normally be joined to a marker for selection in plant cells. Conveniently, the marker may be resistance to a

herbicide, e.g. phosphinothricin or glyphosate, or more particularly an antibiotic, such as kanamycin, G418, bleomycin, hygromycin, chloramphenicol, or the like. The particular marker employed will be one which will allow for selection of transformed cells compared with cells lacking the introduced recombinant DNA.

The fusion peptide in the expression cassette constructed as described above, expresses at least preferentially in developing seeds. Accordingly, transformed plants grown in accordance with conventional ways, are allowed to set seed. See, for example, McCormick et al. (1986, *Plant Cell Reports*, 5: 81-84). Northern blotting can be carried out using an appropriate gene probe with RNA isolated from tissue in which transcription is expected to occur, such as a seed embryo. The size of the transcripts can then be compared with the predicted size for the fusion protein transcript.

Oil body proteins are then isolated from the seed and analyses performed to determine that the fusion peptide has been expressed. Analyses can be for example by SDS-PAGE. The fusion peptide can be detected using an antibody to the oleosin portion of the fusion peptide. The size of the fusion peptide obtained can then be compared with predicted size of the fusion protein.

Two or more generations of transgenic plants may be grown and either crossed or selfed to allow identification of plants and strains with desired phenotypic characteristics including production of recombinant proteins. It may be desirable to ensure homozygosity of the plants, strains or lines producing recombinant proteins to assure continued inheritance of the recombinant trait. Methods of selecting homozygous plants are well know to those skilled in the art of plant breeding and include recurrent selfing and selection and anther and microspore culture. Homozygous plants may also be obtained by transformation of haploid cells or tissues followed by regeneration of haploid plantlets subsequently converted to diploid plants by any number of known means, (e.g.: treatment with

Meth d of Separating Target Molecules Using the Affinity Matrices

As hereinbefore mentioned, the present invention relates to a method of separating a target molecule from a sample using the above described oil body proteins and in some cases, ligands. In the method of the invention, oil bodies are mixed with a sample containing the desired target and the interaction between the ligand and target results in the non covalent association of the target with the oil body. Following centrifugation, the oil bodies and affinity-bound target are separated from the aqueous phase, effectively purifying the target from any contaminants present in the original sample. Repeating the washing step ensures that any remaining contaminants are removed.

Following their attachment to oil bodies, targets may be eluted under conditions determined empirically for each individual ligand-target pair. Treatment of the bound matrix with the appropriate eluent and centrifugation enables recovery of the purified target in the aqueous phase. If the target is a ligand-protein fusion containing a protease recognition site, then it may be treated with the appropriate protease to remove the ligand. The free ligand may then be separated from the target protein by re-application of the oil body affinity matrix or through conventional protein purification methods.

The chemical and physical properties of the affinity matrix may be varied in at least two ways. Firstly, different plant species contain oil bodies with different oil compositions. For example, coconut is rich in lauric oils (C12), while erucic acid oils (C22) are abundantly present in some *Brassica* spp. Furthermore, proteins associated with the oil bodies will vary between species. Secondly, the relative amounts of oils may be modified within a particular plant species by applying breeding and genetic engineering techniques or a combination of these known to the skilled artisan. These techniques aim at altering the relative activities of enzymes controlling the metabolic pathways involved in oil synthesis. Through the application of these techniques, seeds with a sophisticated set of different oils are obtainable. For example, breeding efforts have resulted in the development of a rapeseed with a low erucic acid content (Canola) (Bestor, T. H., 1994, Dev. Genet. 15: 458) and plant lines with oils with alterations in

the position and number of double bonds, variation in fatty acid chain length and the introduction of desirable functional groups have all been generated through genetic engineering (Töpfer et al., 1995, Science, **268**: 681-685). Using similar approaches a person skilled in the art will be able to further expand on the presently available sources of oil bodies. Variant oil compositions will result in variant physical and chemical properties of the oil body fraction. Thus by selecting oilseeds or mixtures thereof from different species or plant lines as a source for oil bodies, a broad repertoire of oil body matrices with different textures and viscosities may be acquired.

Applications of Oil Body Affinity Matrices

Given that it is possible to engineer oil body affinity matrices for several classes of proteins, multiple uses for oil body based affinity matrices are envisioned. Bacteria, fungi, plants and animals all contain proteins which are able to specifically interact with agents such as ions, metals, nucleic acids, sugars, lipids and other proteins. These agents may be immobilized using oil body technology.

The oil body protein affinity matrices can be used to isolate any target molecule that can bind to the oil body protein, either directly or indirectly through a ligand molecule. Examples of target molecules that may be isolated from a sample using the methodology of the present invention include proteins, peptides, organic molecules, lipids, carbohydrates, nucleic acids, cells, cell fragments, viruses and metals. In particular, the inventors have shown that the affinity matrix of the present invention can be used to separate therapeutic proteins (such as thrombin), antibodies, metals (such as cadmium), carbohydrates (such as cellulose), organic molecules (such as biotin) and cells (such as bacterial cells).

Oil body affinity matrices may also be used to separate cells of industrial or medical interest from a mixed population of cells. For example haematopoietic stem cells, which are a subpopulation of blood cells and are used in bone marrow transplantations and in stem cell gene therapies, may be separated from other blood cells using oil body based affinity technology. In recombinant DNA technology it is often required that cells in which recombinant DNA has been successfully introduced, known as transformed cells, are distinguished and separated from cells which failed to

acquire recombinant DNA. Provided that part of the recombinant DNA expresses a cell surface protein which is complementary to a oil body based affinity ligand, it is possible to utilize oil bodies to separate transformed cells from untransformed cells. Oil body affinity technology may also be used to separate cellular organelles such as chloroplasts and mitochondria from other cellular material. Viral particles may also be separated from complex mixtures.

It is also possible to immobilize a class of proteins known as metalloproteins, which contain prosthetic groups that specifically bind ions. Examples of metalloproteins are haemoglobin, which binds iron, parvalbumbin which binds calcium and metallothionein a protein which binds zinc and other metal ions. It is envisioned that oil bodies could be used to scavenge metals from streams of flowing material, which might be water contaminated with the waste of metals from laboratories and industrial processes. Example 4 given below further illustrates this application. Other examples where proteins may be bioimmobilized and employed in a bioremediation strategy include the removal of phosphates, nitrates and phenols from waste streams. In part this approach may overcome the real or perceived limitations of bacterial bioremediation. In certain instances it may not be practical or necessary to rely on affinity partitioning technology to separate the oil body matrix from the target compound. In these instances, it is envisioned that oil bodies may be immobilized on a solid inert surface which could be a flat surface or the surface of a column. A solution containing the affinity ligand may then be passed over the surface coated with immobilized oil bodies whereupon selective affinity binding occurs. It is envisioned that immobilized oil bodies may be used in pipes and in ponds to assist in bioremediation.

Oil body affinity matrices can be used to isolate a recombinant polypeptide from cells. In such a case the recombinant polypeptide (i.e. the target molecule) can associate either directly or indirectly with the oil bodies. This embodiment of the present invention is particularly advantageous as it allows the rapid and inexpensive manufacture of valuable, recombinantly expressed polypeptides.

Accordingly, the present invention provides a method for the isolation of a recombinant polypeptide from a cell, said cell comprising oil bodies and the recombinant polypeptide, said method comprising:

(1) contacting (i) said oil bodies with (ii) said recombinant polypeptide to allow said recombinant polypeptide to associate with said oil bodies; and

(2) isolating said oil bodies associated with said recombinant polypeptide.

In accordance with the present invention, the cell may be any cell comprising oil bodies and a recombinantly expressed polypeptide. Preferably a plurality of cells is used. Suitable cells in accordance with the present invention include any animal cell, plant cell, fungal cell, yeast cell (Leber, R. et al., 1994, Yeast 10: 1421-1428), bacterial cell (Pieper-Furst et al., 1994, J. Bacteriol. 176: 4328-4337) or algae cell (Rossler, P.G., 1988, J. Physiol. (London), 394-400) comprising oil bodies and a recombinantly expressed polypeptide. Preferably however plant cells are used and more preferably plant seed cells.

The recombinant polypeptide may be any polypeptide which is recombinantly expressed by the cell. The polypeptides that may be used in accordance with the present invention may comprise a signal sequence that allows the direction of the polypeptide to a selected sub cellular compartment. Signal sequences that may be used include for example endoplasmatic reticulum retention signals, apoplast targeting sequences, for example the signal sequence from the tobacco pathogenesis relating

Plants: A Production System for Industrial and Pharmaceutical Proteins, John Wiley & Sons Ltd.)

In a preferred embodiment, the recombinant polypeptide target molecule associates with the oil bodies indirectly through a ligand molecule. The ligand molecule may be any molecule capable of associating with an oil body and the recombinant polypeptide. In preferred embodiments one ligand is used, however two or more ligands may be used if desired. Where more than one ligand is used, ligands are selected to associate with each other. The ligand may associate with the oil body and with the recombinant polypeptide through non-covalent interactions, for example by using a bivalent ligand. The ligand may be also be covalently linked to the recombinant polypeptide or to the oil body. In particularly preferred embodiments of the invention, the ligand molecule is a polypeptide. Accordingly, the present invention further provides a method of isolating a recombinant polypeptide from a cell comprising oil bodies, said method comprising:

a) introducing into said cell (i) a first nucleic acid sequence molecule encoding a recombinant polypeptide and (ii) a second nucleic acid sequence encoding a ligand capable of associating with said recombinant polypeptide and with said oil bodies;

b) growing said cell under conditions permitting the expression of said recombinant polypeptide and said ligand;

c) contacting (i) said oil bodies with (ii) said recombinant polypeptide to allow said recombinant polypeptide to associate with said oil bodies through said ligand; and

d) isolating said oil bodies associated with said recombinant polypeptide.

In embodiments where the ligand is a polypeptide, the ligand polypeptide may conveniently be prepared as a fusion protein with the target recombinant polypeptide and recombinantly expressed in the cell. In such an embodiment, the ligand is any molecule that can bind to, or associate with, the oil bodies or oil body protein but is preferably not a protein that is normally associated with the oil bodies. The term "a protein that is normally associated with the oil bodies" includes proteins that are

normally associated with oil bodies in non-transformed or normal cells such as oil body proteins (for example oleosins) or proteins that are naturally present in normal or non-transformed cells and may associate with oil bodies when oil bodies are purified.

In order to prepare the fusion protein, a chimeric nucleic acid sequence is prepared that comprises (a) a nucleic acid sequence encoding the recombinant polypeptide linked to (b) a nucleic acid sequence encoding the ligand. The fusion protein may comprise a cleavage site, for example a chemical or enzymatic cleavage site, that allows for the separation of the target from the ligand molecule. Preferred protein ligand molecules in accordance with the present embodiment include antibodies and fragments thereof (i.e. Fab, $F(ab^{1})_{2}$, monoclonal antibodies, single chain antibodies, recombinantly produced binding partners). However, any protein or peptide capable of associating with the oil body and the recombinant polypeptide may be used.

By "contacting the oil bodies with the recombinant polypeptide", it is meant that the oil bodies are brought into proximity of the recombinant polypeptide in a manner that allows the recombinant polypeptide to associate with the oil bodies. In one embodiment, contacting of the recombinant polypeptide and the oil body is accomplished following the application of a technique resulting in the substantial disruption the cell's integrity. Generally any technique that substantially releases the cell's constituents may be used, however the technique typically varies depending on the cell type that is selected. Techniques to disrupt cells include physical techniques such as the application of high pressure, as well chemical and biochemical techniques such as the use of enzymes capable of degrading cellular membranes and other art-recognized techniques. Preferably the techniques and conditions to substantially disrupt the cell's integrity are selected such that the cell's constituents are released while the oil bodies remain substantially intact. In embodiments of th invention where plant seeds are used, grinding equipment such as mills, for example colloid mills,

invention in which the recombinant polypeptide associates with the oil body after substantially disrupting the cell is particularly desirable because it allows expression of the recombinant polypeptide in a sub-cellular compartment, for example the golgi complex, endoplasmatic reticulum or apoplast, having the most appropriate physico-chemical conditions for the selected recombinant polypeptide while taking advantage of the oil body matrix as an extremely cost-effective purification tool.

In another embodiment of the invention, contacting of the recombinant polypeptide and oil bodies is accomplished within the cell. In such an embodiment the recombinant polypeptide is expressed in a manner that allows the recombinant polypeptide to be directed intracellularly to the oil bodies. This could involve the expression of the recombinant polypeptide in a manner that allows the polypeptide to accumulate in the cytoplasm.

By the term "isolating" it is meant that the oil bodies associated with the recombinant polypeptide are separated from other cellular constituents. The degree of purity may vary and generally depends on the desired purity of the recombinant polypeptide. In general, separation of the oil bodies may be performed as hereinbefore described. The recombinant polypeptide may further be separated from the oil body as hereinbefore described. In embodiments of the invention where the recombinant polypeptide is fused to a protein-ligand and the fusion protein comprises a cleavage site, the recombinant polypeptide and the ligand may be separated by performing a cleavage reaction. Such a reaction may be performed while the fusion protein is associated with the oil body or upon prior separation of the fusion protein from the oil body. An essentially pure recombinant polypeptide may be obtained using additional purification tools such as for example column chromatography.

The following examples illustrate various systems in which oil bodies can be used as affinity matrices. It is understood that the examples given below are intended to be illustrative rather than limiting.

EXAMPLES

EXAMPLE 1

Purification of Thrombin

The following example demonstrates the utility of an oil body affinity matrix for the purification of thrombin. Thrombin is a serine protease which plays a central role in blood coagulation. It cleaves fibrinogen to produce fibrin monomers which polymerize to form the basis of a blood clot (Fenton 1981; *Ann. N.Y. Acad. Sci.* **370**: 468-495). Alfa-thrombin consists of two polypeptide chains of 36 (A-chain) and 259 (B-chain) residues linked by a disulphide bridge. Degen *et al.* 1983; *Biochemistry* **22**: 2087-2097). Hirudin, which is found in the salivary glands of the medicinal leech *Hirudo medicinalis*, is a very specific and potent inhibitor of thrombin. This inhibition is a result of the non-covalent binding of hirudin to specific parts of the alfa-thrombin chain. (Stone and Hofsteenge 1986; *Biochemistry* 25: 4622-4628).

The immobilized ligand is comprised of an isoform of hirudin fused to the 18 kDa *Arabidopsis* oleosin (oil body protein) (Van Rooijen *et al.*, 1992; *Plant Mol. Biol.* 18: 1177-1179). Expression of the construct is regulated by the Arabidopsis 18 kDa oleosin promoter (Van Rooijen *et al.*, 1994; *Plant Mol. Biol.* 18: 1177-1179). The sequence of the oleosin-hirudin fusion is shown in Figure 2 and in SEQ.ID.NO:3.

Oleosin-Hirudin Construct

Oligonucleotide primers were designed based upon the reported sequence for a *Brassica napus* oleosin gene (Murphy *et al.* 1991, *Biochim. Biophys. Acta* **1088**: 86-94) and used to amplify a fragment from *B. napus* genomic DNA through PCR. Using this fragment as a probe, a clone carrying a 15 kbp insert was identified and isolated from a EMBL3 *Arabidopsis* genomic library. Oligonucleotide primers were used to amplify a fragment from this insert containing the entire oleosin coding sequence and intron together with 840 basepairs of the 5′ upstream region. The primers were designed so as to eliminate the translational stop codon and to introduce a *PstI* restriction endonuclease recognition site at the 5′ end and a *SalI* followed by a *PvuI* site at the 3′ end of the fragment. The fragment was end-filled and ligated into the *SmaI* site of the plasmid vector pUC19. A *SalI*

- *EcoRI* fragment from plasmid pBI121 (Clontech) comprising the nopaline synthetase terminator sequence was then inserted to generate pOBILT.

A synthetic hirudin variant 2 (HV2) sequence was synthesized based upon reported sequence information (Harvey *et al.* 1986, *Proc. Natl. Acad. Sci. USA* 83: 1084-1088) but employing *B. napus* and *Arabidopsis* codon usage. The sequence was amplified using four overlapping oligonucleotide primers designed such that the resulting fragment possessed *PvuI* and *SalI* sites at the 5′ and 3′ ends respectively. This fragment was ligated into the *SmaI* site of the pUC19 plasmid vector to generate pHIR. The *PvuI* - *SalI* fragment from pHIR was then inserted into pUCOBILT between the oleosin and terminator sequences to form an in-frame fusion with the oleosin coding region giving pUCOBHIRT. The entire construct was subcloned into pBluescript KS+ (pBIOBHIRT) and then into the *PstI* site of pCGN1559 plasmid (McBride and Summerfelt, 1990, *Plant Mol. Biol.* 14: 269-276) carrying a neomycin phosphotransferase gene under control of the 35-S CaMV promoter (pCGOBHIRT). This plasmid was introduced into *Agrobacterium tumefaciens*. The preparation of this plasmid is shown in Figure 3.

Transformation and Regeneration

Procedures for the transformation of *Agrobacterium* and plants have been described previously. *Agrobacterium tumefaciens* was transformed with the above construct through electroporation (Dower et al., 1988; *Nucl. Acids Res.* 16: 6127-6145). The transformed bacteria were then used to transform cotyledonary explants of *Brassica napus*, followed by plant regeneration according to the methods of Moloney *et al.* (1989; *Plant Cell Reports* 8: 238-242). Transgenic plant were initially identified using a neomycin phosphotransferase assay and subsequently confirmed by expression of the oleosin-hirudin fusion as determined through northern and immunoblot analysis.

Preparation of Oil Bodies

Seed from either control (non-transgenic) plants or transgenic plants expressing the oleosin-hirudin fusion were homogenized in five volumes of cold grinding buffer (50 mM Tris-HCl, pH 7.5, 0.4 M sucrose and 0.5 M NaCl) using a polytron operating at high-speed. The homogenate was centrifuged at approximately 10 x g for 30 min. to remove particulate

matter and to separate oil bodies from the aqueous phase containing the bulk of soluble seed protein. Oil bodies were skimmed from the surface of the supernatant with a metal spatula and placed in one volume of fresh grinding buffer. To achieve efficient washing in subsequent steps, it was important to ensure that the oil bodies were thoroughly redispersed. This was accomplished by gently re-homogenising the oil bodies in grinding buffer with the polytron operating at low-speed. Using a syringe, the resuspended oil bodies were carefully layered underneath five volumes of cold 50 mM Tris-HCl, pH 7.5 and centrifuged as above. Following centrifugation, the oil bodies were again removed and the washing procedure repeated three times to remove residual contaminating soluble seed proteins. The final washed oil body preparation was resuspended in one volume of cold 50 mM Tris-HCl pH 7.5, redispersed with the polytron, and was then ready for use as an affinity matrix.

Affinity Purification of Thrombin

The purification of thrombin using the oleosin-hirudin fusion protein is shown schematically in Figure 4. In order to evaluate the binding of thrombin, affinity matrices were prepared from transgenic *Brassica napus* seeds expressing the oleosin hirudin fusion protein (4A4 seeds) (Parmenter *et al. Plant Molecular Biology* (1995) **29**: 1167-1180) and from wild type *Brassica napus* cv Westar seeds. Binding of thrombin to both matrices was evaluated. Procedures for the preparation of washed oil bodies from seeds were the same as those described above. Solutions containing a range of thrombin activities between 0 and 1 units were mixed with 10 µl of a fixed amount of affinity matrix (prepared from a total of 10 mg of dried seeds; corresponding to approximately 100 µg of total oil body protein) in 500 µl binding buffer (50 mM Tris-HCl (pH 7.5); 0.1% (w/v) BSA). The oil body suspension was then incubated for 30 minutes on ice and centrifuged at 14,000 rpm for 15 minutes at 4°C. The buffer under the oil bodies (termed 'unternatant') containing the unbound, free thrombin was recovered using an hypodermic needle and assayed for thrombin activity as follows. A total of 250 µl of unternatant was added to 700 µl binding buffer and prewarmed to 37°C. Following the addition of 50 µl of 1 mM thrombin substrate N-p-tosyl-gly-pro-arg-p-nitroanilide (Sigma) to the unternatant, the change

in optical density at 405 nanometers was monitored spectrophotometrically for 3 minutes. The concentration of thrombin in the assay mixture was determined employing a standard curve which was constructed using a set of thrombin samples containing known concentrations of thrombin. The values obtained from these assays were used to calculate the concentration bound thrombin assuming:

[bound thrombin] = [total thrombin] - [free thrombin]

The ratio of the concentration of bound over the concentration of free thrombin was plotted as a function of the concentration of bound thrombin (Scatchard plot). From these plots the dissociation constants of the affinity matrix were calculated following standard procedures (Scatchard, G. *Ann. N.Y. Acad. Sci.* (1949) 57: 660-672) and assuming: $K_a = 1/K_d$. The dissociation constants of the affinity matrices were 3.22×10^{-7}m for wild type and 2.60×10^{-8}m for 4A4 oil bodies.

In order to evaluate the recovery of bound thrombin from the matrices a NaCl gradient was employed. The elution profile of thrombin bound to oleosin-hirudin oil body matrices was compared with the profile from thrombin bound to wildtype oil body matrices. Procedures for preparation of wild type oil bodies from wild type *Brassica napus* cv Westar seeds and for the preparation of oleosin-hirudin oil bodies from *Brassica napus* 4A4 seeds (Parmenter et al. *Plant Molecular Biology* (1995) 29: 1167-1180) were identical to those described above. Procedures for binding of thrombin to the matrices were as described above, except 100 µl aliquots of oil bodies were used to bind 0.5 units of thrombin. Oil body suspensions were left on ice for 30 minutes prior to centrifugation for 15 minutes at 4°C and 14,000 rpm. The supernatant was assayed for (unbound) thrombin activity. The oil body matrix was then resuspended in binding buffer to which NaCl was added to a final concentration of 0.05 M. Starting with the 30 minutes incubation of the oil body suspension on ice, the procedure was repeated five times increasing the NaCl concentration in a stepwise fashion. The final NaCl concentrations used were 0.05 M, 0.1 M, 0.2 M, 0.3 M, 0.4 M and 0.6 M. The NaCl concentrations in the thrombin assay were kept constant at 150 mM. Figure 5 shows the elution profiles obtained when wildtype oil bodies and 4A4 oil bodies were used.

EXAMPLE 2

Use of Antib dies as Bivalent Ligands

Antibodies may be used as bivalent ligands by virtue of their affinity both for specific epitopes and for other antibodies or proteins (for example the *Staphylococcus aureus* protein A) which have affinity for immunoglobulins (IgGs). In this example, polyclonal anti-oleosin antibodies serve as a bivalent ligand and antibodies raised in rabbits against the anti-oleosin antibodies serve as the target. This example is illustrated schematically in Figure 6.

Oil bodies were prepared from 5 g of wild type *Brassica napus* cv Westar seeds following the procedure described in Example 1. Subsequently, oil bodies were washed twice with 100 mM glycine (pH 2.5), neutralized through two washes in binding buffer (50 mM Tris-HCl, pH 7.5) and resuspended in 5 ml of binding buffer. A 150 µl aliquot of the washed oil body preparation was combined with 500 µl of rabbit serum containing anti-oleosin antibodies (ligand antibodies), diluted 1:10 with binding buffer. The oil body suspension was mixed thoroughly and incubated for 1 h at 4°C with agitation. Following incubation, unbound ligand antibodies were removed from the oil body suspension through three washes with 1 ml of binding buffer. Oil bodies were then combined with 500 µl of serum diluted 1:500 in binding buffer and containing anti-rabbit IgG antibodies (the target antibodies) conjugated with horseradish peroxidase (HRP) as a detection label (Sigma). This suspension was mixed and incubated under conditions identical to those used for the anti-oleosin antibody binding. As a control, target antibodies were incubated with oil bodies which had not been previously bound to ligand antibodies. Both samples were subsequently washed four times with 1 ml of binding buffer to remove unbound antibodies. Using binding buffer, the samples were equalized with respect to concentration of oil bodies as determined by measuring sample turbidity spectrophotometrically at 600 nm. To assay for bound target antibody, samples containing 5 µl of oil bodies were mixed with 1 ml of the HRP colorimetric substrate tetramethylbenzidine in 0.01% hydrogen peroxide and reacted for 10 minutes at room temperature. Reactions were stopped by the addition of 500 µl of 1 M H_2SO_4 and the absorbance at 450 nm was

determined. Corrections for the presence of residual, unbound target antibody remaining after washing were made by assaying 5 μl of the final wash fraction. The results obtained for control and ligand bound oil body preparations are set forth in Figure 7.

EXAMPLE 3

Use of Oleosin-Specific Ligands

The use of an oleosin-specific ligand represents an alternative to the use of an antibody or genetically-engineered oleosin fusion proteins for the purification of recombinant target proteins. In this case, the target protein is fused to the oleosin-specific ligand and the endogenous oleosins present on the oil bodies of non-transgenic seeds serve as the complementary ligand-affinity matrix. In addition to eliminating the requirement for a transgenic line expressing an oleosin fusion, this approach increases the overall capacity of the affinity matrix, since all of the endogenous oleosins may now participate in binding.

Oleosin-specific ligands may be identified and isolated from a peptide phage display library screened with oleosin protein. Since the extreme hydrophobicity of the oleosin central domain can result in aggregation and precipitation of the protein when removed from oil bodies, a mutant protein lacking this domain may be used for screening. This has little effect on the efficacy of the ligand, as only the hydrophillic portions of the oleosin are exposed to the cytoplasm (i.e. the N- and C-termini). Hence, these are the only regions available for binding to a ligand. Once isolated, the ligand may be fused to a common reporter protein, green fluorescent protein (GFP) (Prasher, 1995, *Trends Genet.* 11:320 323), to demonstrate purification.

Removal of the Oleosin Central Domain

Oligonucleotide primers specific for the *Arabidopsis* oleosin gene described above can be used to amplify an oleosin gene from a *B. napus* cDNA library (van Rooijen 1993, *Ph.D. Thesis,* University of Calgary). Primers flanking sequences encoding the N-terminal 62 amino acids and the C-t rminal 55 amino acids, may be used to amplify sequences for the respective N- and C-terminal oleosin domains in separate reactions. Additionally, the primer for the 5′ end of the N-terminal domain contains a

sequence for a thrombin recognition site to enable cleavage of the fusion protein as described below. The resulting fragment was ligated into the *SmaI* site of the bacterial expression vector pEZZ 18 (Pharmacia). This vector contains sequences encoding a signal peptide for protein secretion into the periplasm, and synthetic IgG binding domains derived from protein A to facilitate protein purification, downstream of the multiple cloning site.

Expression and Purification of the Oleosin Deletion Construct

The vector carrying the deletion mutant construct is introduced into *E. coli* using standard methods and transformants selected. A culture of the transformed bacteria can be induced to express the synthetic protein A-mutant oleosin fusion protein by addition of 1 mM IPTG. Induced cells may be pelleted and resuspended in 5 mM $MgSO_4$ causing lysis of the periplasmic membrane through osmotic shock. The lysed cells are centrifuged and the supernatant containing the secreted protein is loaded on to a column containing IgG-coupled sepharose. After washing to remove unbound protein, the column is loaded with a buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl and 1.0U/ml of purified Bovine thrombin (Sigma) to cleave the mutant oleosin from the synthetic protein A. Following incubation at 37°C for 4h, the column is drained and the eluate passed through a column of heparin-coupled sepharose to remove thrombin. The eluate from this column, containing the mutant oleosin protein, is recovered and purity of the protein examined through gel electrophoresis followed by staining with Coomassie blue R250.

Generation of a Peptide Combinatorial Library

A random peptide combinatorial library may be generated according to the methods of Scott and Smith (1990; *Science* 249: 386-390). Briefly, the PCR is used to amplify a synthetic DNA fragment containing the degenerate sequence $(NNK)_6$; where 'N' represents an equal mixture of deoxynucleotides G, A, T, and C, and K represents an equal mixture of deoxynucleotides G and T. The degenerate sequence encodes for hexameric peptides among which are represented every possible combination of the 20 amino acids and amber stop codon. The PCR product is ligated into the gene III sequence of the filamentous bacteriophage fUSE and the resulting [illegible] through electroporation.

Identification and Isolati n of Oleosin-Specific Ligands

The peptide phage display libraries are amplified, concentrated and stored in aliquots of 10^{12} tdu/ml. Purified mutant oleosin protein is biotinylated using a thiol-cleavable linker (S-S biotin, Pierce) and purified by size exclusion chromatography. Aliquots of the peptide phage display library containing 5×10^{11} tdu in two ml are screened with the biotinylated protein at a concentration of 50 nM. Phage binding the mutant oleosin protein are recovered using streptavidin-coated paramagnetic beads. Following washing, the phage are eluted through the addition of 50mM dithiothreitol which cleaves the disulphide bond. The eluted phage are then incubated with an excess of log-phase F+ *E. coli*. Aliquots of the infected cells are plated to determine the phage titre and the remaining cells used in successive rounds of amplification and screening. Following enrichment of the eluted phage by 3-4 orders of magnitude, individual phage are selected and tested for binding to mutant oleosin by direct ELISA. Binding by phage is detected using anti-phage antibodies (Crosby and Schorr, 1995, In *Annual Review of Cell Biology*). Single stranded DNA is isolated from phage exhibiting binding and the peptide-encoding sequence determined.

Affinity Purification with Oleosin-Specific Ligands

The sequence for an oleosin ligand isolated as described above is fused in-frame upstream the sequence for *gfp10* (Prasher *et al.*, 1992, *Gene* 111: 229-233) encoding GFP and the construct ligated into the bacterial expression vector pKK233 (Pharmacia). Soluble protein is extracted through sonication of cells induced to express the ligand-GFP fusion, and adjusted to a concentration of 10 mg/ml in 50 mM Tris-HCl, pH 7.5.

Twenty ml of the protein solution is mixed with 2ml of oil bodies prepared as described above, from seeds of non-transgenic plants. The mixture is incubated at 4°C for 30 min with agitation to allow binding and then centrifuged to separate the oil bodies and soluble fraction. The amount of GFP remaining in the soluble fraction after removal of oil bodies is determined by fluorescence spectrofluorometry at a wavelength of 508 nm and compared with that in the original bacterial extract. The amount of bound GFP is calculated to determine the capacity of the matrix.

amount of cells present in the cell pellet as the concentration of anti-IgGs present in the oil-body *S. aureus* mixture increases.

<u>Differential Binding of Two Strains of *Staphylococcus aureus*.</u>

In this experiment an oil body affinity matrix is employed to demonstrate differential binding of two strains of *Staphylococcus aureus*. Formalin fixed *S. aureus* strains, one expressing the IgG binding surface antigen protein A and one lacking protein A, are commercially available from Sigma. Dilute aliquots of both *S. aureus* strains of equal OD_{550} could be prepared. To each of these aliquots control oil bodies from untransformed plants or oil bodies mixed with anti-oleosin antibodies could be added. Following incubation for an appropriate length of time at an appropriate temperature, the samples could be centrifuged to pellet unbound bacterial cells and to separate the oil body fraction. The oil bodies could be decanted, vortexed and the OD_{550} could be determined. The pellets could be resuspended and the OD_{550} of the unternatant could be determined. It is anticipated that only in the sample containing the *S. aureus* strain expressing protein A and the oil body complexed with anti-oleosin antibodies, fractionation of these cells to the oil body fraction will be observed. Binding of the cells to the oil body could be further demonstrated by lowering of the pH of the oil body fraction. Subsequent to centrifugation the release of cells from the oil bodies could be evidenced by the presence of a pellet and/or an increase in OD_{550} upon resuspension of the pellet.

<u>Separation of *Staphylococcus aureus* from *E. coli*</u>

A viable *S. aureus* strain could be mixed with varying quantities of cells of an *E. coli* strain having a specific antibiotic resistance. The mixed bacterial sample could be vortexed with control antibodies and with oil bodies which have been complexed with anti-oleosin antibodies. After incubation for an appropriate length of time and at an appropriate temperature oil bodies could be washed and the unternatant and oil bodies could be directly titrated and selectively plated on blood agar for *S. aureus* growth and on LB plates for *E. coli* growth. The enrichment or actual separation obtained could be determine by an estimate of colony forming

EXAMPLE 5

Separation of Whole Cells

The following example illustrates the capacity of oil bodies to immobilize whole cells. One potential for the use of bacterial cell separation lies in the utility for diagnostics. It is also desirable to separate unique eukaryotic cells such as lymphocytes and stem cells from complex mixtures of cells where the cell type of interest is present in relatively low numbers.

Binding of Staphylococcus aureus to oil bodies via protein A

For the purpose of this example, *S. aureus* cells, which express protein A as a surface antigen were mixed with oil bodies with varying amounts of polyclonal anti-oleosin antibodies.

Preparation of oil bodies

Seeds of *B. napus* cv Westar were surface sterilized in bleach, rinsed and ground with a mortar and pestle in grinding buffer (50 mM Tris pH 7.5, 0.4 M sucrose and 100 mM glycine). The homogenate was filtered through Miracloth into sterile 15 ml Corex tubes. The filtered homogenate was then centrifuged for at 4°C for 10 min at 10,000 xg. The oil body fraction was removed and resuspended in 50 mM Tris pH 7.5 and 0.4 M sucrose and washed two times using the same buffer. Aliquots of 1 ml oil bodies were transferred to 1.5 ml Eppendorf tubes and centrifuged at room temperature for 10 min at 16,000 xg. The oil bodies were washed in 50 mM Tris pH 7.5 and 0.4 M sucrose 5-6 more times until no visible pellet was observed.

Binding of S. aureus cells to anti-oleosin coated oil bodies

Formalin fixed *S. aureus* cells (Sigma, P-7155) were washed 3-4 times in 50 mM Tris-Cl pH 7.5. and resuspended. Washed oil bodies (300 µl) and *S. aureus cells* (were mixed with varying amounts of anti-oleosin IgGs (50 µl). After mixing and incubating at room temperature for 2 hrs, the mixtures were centrifuged at room temp at 16,000 xg for 5 min. The oil body fraction and unternatant were carefully removed and the cell pellet was washed twice in 1 ml 50 mM Tris-Cl pH 7.5. The walls of the tube were wiped with a tissue to remove traces of oil. Subsequently the drained cell pellets were resuspended in 1 ml of water and the OD_{600} were determined.

was digested with *KpnI* to release the oleP-oleMT-ubi3' insert. This expression cassette was inserted at the *KpnI* site of the binary vector pCCN1559 to yield the final construct pBIOOM3'. The sequence of the oleosin-metallothionein fusion is shown in Figure 8 and SEQ.ID.NO.6. The construction of plasmid pB100M3' is shown in Figure 9.

<u>Transformation and Regeneration</u>

Transgenic *B. carinata* plants expressing the oleosin-metallothionein fusion were created using transformation and regeneration protocols as described in Example 1.

<u>Oil Body Preparation</u>

Washed oil bodies were prepared from *B. carinata* seeds of transgenic and control plants as described in Example 1.

<u>Removal of Cd++ From Solution Using an Oil Body Affinity Matrix</u>

The use of the oleosin-metallothionein fusion to bind cadmium ions in solution is shown schematically in Figure 10.

A solution of 10 µM $CdCl_2$ in 10 mM Tris-HCl, pH 7.2 containing 0.01 µCi/ml ^{109}Cd was prepared. A 1 ml aliquot of this $CdCl_2$ solution was thoroughly mixed with 100 µl of washed oil bodies (1.6 mg oil body protein) prepared from seeds expressing the oleosin-metallothionein fusion protein and incubated at 22°C for 1 hr. Following centrifugation for 5' at 10,000 xg to separate the oil bodies from the aqueous phase and 2 washes in 1 ml of 10 mM Tris-Cl, pH 7.2, the amount of $^{109}Cd^{++}$ remaining bound to oil body fraction was determined using a gamma-counter (Cobra auto-gamma, Canberra Packard, Canada). An identical experiment was performed with oil bodies from non-transgenic seeds to detect and correct for non-specific binding of Cd ions to the matrix.

Cd^{++} ions were eluted from the oil body metallothionein affinity matrix by mixing of the oil body fraction with 1 ml of 100 mM

The oil bodies are washed twice in 20 ml of 50 mM Tris-HCl, pH 7.5, resuspended in 2 ml of the same buffer and divided into 20 aliquots of 100 µl. Conditions for the elution of ligand-GFP fusion protein are determined by adding 1ml of solutions ranging in pH from 2-10 and in NaCl concentration from 0-1 M to different aliquots. After mixing and incubation at 4°C for 30 min, the oil bodies are removed and the soluble fractions collected. The amount of ligand-GFP fusion protein in the soluble fraction is determined by fluorescence spectrophotometry.

<u>EXAMPLE 4</u>

Removal of Heavy Metal Ions

The following example demonstrates the utility of oil body affinity matrices for the recovery/removal of non-protein targets from complex solutions. For the purpose of this example the metallothionein/Cd^{++} ligand pair was used. However other metal binding proteins such as phytochelatins (Rauser, 1990; *Ann. Rev. Biochem*; 59: 61-86) and metal ions including Cu^{++} and Zn^{++} could also be used.

<u>Oleosin-Metallothionein Fusion</u>

An oleosin gene from a *B. napus* cDNA library (van Rooijen 1993, *Ph.D. Thesis*, University of Calgary) was amplified through PCR with oligonucleotide primers designed so as to create *NotI* and *NcoI* sites at the 5′ and 3′ ends of the gene respectively. The resulting fragment was digested and placed into the *NotI/NcoI* sites of pCN to yield plasmid poleGN. The human metallothionein gene, *mt-II* (Varshney and Gedamu, 1984, *Gene*, 31: 135-145) was amplified using oligonucleotide primers designed to create a unique *NotI* site at the 3'-end of the gene. The resulting PCR product was subcloned into the blunt-end *EcoRV* site of pBluescript KS+ to form pBSMTC. The *mt-II* gene was then excised from this plasmid and subcloned into the *NcoI/KpnI* sites of poleGN replacing the GUS-NOS region to generate pOLEMTC. The 773 base oleosin-MT fusion of pOLEMTC was excised with *NotI* digestion and inserted into the unique *NotI* site of polePN3' between the oleosin promoter (oleP; Van Rooijen *et al.*, 1992, *Plant Mol.*

Identification of Pathogens Present in Low Concentrations in a Complex Mixture

For diagnostic purposes it is often desirable to concentrate bacterial or viral pathogens which invade human or animal tissues in low numbers. An oil body affinity matrix could be used to enrich for these pathogens, so that they could subsequently be identified and characterized.

Pathogens often specifically bind to human or animal cells through the interaction with a receptor or surface protein. Oleosin could be fused to the human or animal protein ligand and recombinant oil bodies could be employed to immobilize the pathogens. Examples of the formation of protein complexes formed between proteins of human and pathogenic origins known to the prior art include: human fibrinogen or fibrin specific domains which bind to *S. aureus* protein clumping factor A (clf-A) (McDevitt *et al.* 1995; *Mol. Microbiol.* 16; 895-907); human decay accelerating factor (DAF) to which urinary and intestinal tract pathogenic *E. Coli* bind (Nowicki *et al.* 1993: *J. Of Experim. Med.* **178**: 2115-2121); a human cell ligand which is expressed in the carcinoma cell line Caco-2 and which binds uniquely to the 28 kD *Klebsiella pneumoniae* fimbria protein KPF-28 (Di Maretino *et al.*, 1996; Infect. and Immun. **64**: 2263-2266) and human cell extracellular matrix fibronectin specific domains which complex specifically with *Streptococcus pyrogenes* adhesin (protein F) (Ozeri *et al.*, 1996; EMBO J. **15**: 989-998).

EXAMPLE 6

Separation of Small Organic Molecules

This example describes how an oil body affinity matrix may be used for the recovery/removal of small organic molecules from solution. By way of example, the small organic molecule, biotin, is purified using avidin as a ligand.

Construction of Avidin Ligands

Avidin is a protein synthesized by avian species and exhibits an extremely high affinity for biotin, a natural co-factor for many carboxylases. Preparations of purified avidin (commercially available from Sigma) can be conjugated chemically to anti-oleosin antibodies using standard procedures known to those skilled in the art. This approach would

yield a bivalent avidin ligand suitable to demonstrate affinity based removal of biotin. Alternatively, an oleosin-avidin gene fusion may be utilized. The gene encoding avidin in chicken (*Gallus gallus*) has been identified and its sequence has been determined (Beattie *et al*., 1987, *Nucl Acids Res.* **15**: 3595-3606). Based on the sequence the gene for avidin could be synthesized chemically or through the PCR and fused to the *B. napus* oleosin (van Rooijen, 1993, *Ph.D. Thesis*, University of Calgary) as described in example 4. Streptavidin, an analogous bacterial biotin binding protein, could also be employed.

<u>**Oil Body Preparation**</u>

Washed oil bodies would be prepared from seeds of transgenic plants and/or control plants as described in example 1.

<u>**Binding of Bivalent Avidin-Oleosin Ligand**</u>

Binding of anti-oleosin antibodies and removal of unbound ligand will be as detailed in example 3.

<u>**Removal of Biotin from Solution**</u>

Solutions containing known concentrations of biotin could be combined with a fixed amount of oil bodies complexed with an anti oleosin antibodies conjugated with avidin. Following binding, the mixture would be centrifuged to separate oil body and aqueous fraction. The amount of biotin remaining in the aqueous fraction is determined by competitive ELISA using anti-biotin antibodies conjugated to horse radish peroxidase (HRP). The amount of bound biotin may be calculated assuming:

[bound biotin] = [total biotin] - [free biotin]

From the obtained values, the dissociation constants can be determined as described in example 2. As a control, an identical experiment could be performed with oil bodies bound to anti oleosin antibodies which have not been conjugated with avidin. If desired, biotin could be released from the oil body-avidin matrix through competitive elution using an excess of 2-(4'-hydroxybenzene) benzoic acid (HABA). Elution may also be aided by employing a genetically engineered mutant of avidin which exhibits a lower affinity for biotin. Such mutants have been described for the analogous biotin binding protein from bacteria, streptavidin (Chilkoti *et al*., 1995; Bio/Technol. 13: 1198-1204).

EXAMPLE 7

Separation of Carbohydrates

The following example describes the utility of oil body matrices for the recovery of carbohydrates from complex biological mixtures. In this example the inventors demonstrate that an oil body immobilized cellulase is capable of binding cellulose.

Oleosin-Cellulose Binding Domain Fusion

Several of the cellulases produced by the bacterium *Cellulomonas fimi* contain discrete cellulose binding domains (CBDs). These CBDs independently bind to cellulose even when they are separated by proteolytic cleavage or genetic manipulation from the catalytic domain of the enzyme. Plasmid pUC18-CBDPT contains a fragment coding for the CBD of the beta-1,4-glucanase (Gilkes *et al.*, 1992, *Journal of Biol. Chem.* **267**: 6743-6749) and could be used to construct an oleosin-CBD gene fusion. A DNA fragment encoding the CBD domain could be isolated from pUC18-CBDPT using appropriate restriction enzymes or using the PCR. Alternatively, the CBDs of other cellulases from *C. Fimi* or cellulases from other sources could be used. An oleosin gene from *B. Nupus* isolated from a cDNA library (van Rooijen, 1993, *Ph.D. Thesis*, University of Calgary) was cloned in pCN using the PCR and yielding plasmid pOLEGN as described in example 4. An in-frame gene fusion between the oleosin gene and the CBD gene could be generated using standard molecular techniques known to those skilled in the art. The final construct would comprise the CBD domain translationally fused immediately downstream of the oleosin.

Transformation and Regeneration

In order to introduce the fusion gene construct in plants, it would be subcloned in a binary vector, such as pCGN1559. Transgenic plants which express the oleosin-CBD fusion could be generated as described in example 1.

Oil B dy Preparation

Washed oil bodies could be prepared from the seeds of

<u>Removal of Cellulose from Solution Using an Oil Body Affinity Matrix</u>

In order to evaluate binding of cellulose to the oil body affinity matrix, the binding capacities of oil bodies of wild type and transgenic plants are compared. Oil bodies could be mixed with appropriately buffered solutions containing a range of cellulose concentrations. The oil body suspension could then be incubated for an appropriate length of time and at an appropriate temperature. Upon centrifugation, the supernatant could be recovered and assayed for cellulose concentrations. The concentrations bound cellulose and free cellulose could be calculated assuming:

[bound cellulose] = [total cellulose] - [free cellulose]

The ratio of the concentration bound over the concentration free cellulose could be plotted as a function of the concentration of bound cellulose. From these plots dissociation constants could be calculated following standard procedures (Scatchard, G. *Ann. N. Y. Acad. Sci.* (1949) 57: 660-672) and as detailed in example 2.

<u>EXAMPLE 8</u>

Separation of Nucleic Acids

The following example describes a method in which oil bodies are employed to bind single stranded (SS) nucleic acids.

<u>Isolation of Single Stranded Nucleic Acids</u>

A method for capturing SS nucleic acids may be used in diagnostics, such as plant viral disease, or in research applications where non-reannealed SS nucleic acids need to be selectively removed from solutions such as in hybridization reactions for differential screening of expressed genes. Oleosins could be fused with SS DNA or RNA binding proteins or specific domains thereof and could be used to trap SS nucleic acids for identification or further amplification. Well characterized SS nucleic acid binding proteins include: Agrobacterial Ti plasmid Vir E2 protein

74: 1141-1148) and *E. Coli* RecA and single stranded binding (SSB) proteins (Radding, 1991 *J. Biol. Chem.* **266**: 5355-5358).

EXAMPLE 9

Separation of Recombinant Proteins

The following example further demonstrates the utility of an oil body affinity matrix for the purification of recombinant target proteins. For the purpose of this example, the IgG/protein A ligand pair has been chosen. The construct employed consists of a protein A domain which was fused to the 18 kDa *Arabidopsis* oleosin (Van Rooijen *et al.*, 1992; *Plant Mol. Biol.* 18: 1177-1179). Oil bodies containing oleosin-protein A fusion proteins were isolated and used to demonstrate specific binding of rabbit-anti-mouse IgGs conjugated to Horse Raddish Peroxidase (HRP). The configuration of the oleosin-protein A fusion on the oil body and binding of IgG to the fusion is shown in Figure 15.

The Oleosin-Protein A Fusion

A synthetic protein A sequence encoding a protein capable of binding to IgG was synthesized based on reported sequence information (pRIT2T, protein A gene fusion vector; Pharmacia) and was amplified through the PCR. Each primer used in the PCR contained restriction sites 5′ to the protein A-specific sequence in order to facilitate cloning. The reverse primer (i.e. the primer in the antisense direction) also contained a translational stop codon following the coding sequence. Fig 13 shows the position of the PCR primers relative to the protein A sequence. (The protein A sequence and the primer sequences are also separately shown in SEQ.ID.NO:8, SEQ.ID.NO:10 and SEQ.ID.NO:11 respectively). The resulting fragment was ligated into a pUC19 plasmid carrying the *Arabidopsis* oleosin gene comprised of an 867 bp upstream promoter region followed by the coding region (with its associated intron) from which the translational stop codon had been removed. The 3′ end of the construct contains the nopaline synthase transcriptional terminator. A spacer sequence encoding a recognition sequence for the endoprotease thrombin was incorporated immediately downstream of the oleosin coding sequence. The protein A gene sequence was introduced between this spacer sequence and the terminator sequence. In the final expression construct the oleosin and

protein A coding regions were fused in the same reading frame. The entire construct (Figure 14 and SEQ.ID.NO:12) was then excised from the pUC19 plasmid and subcloned into the plant transformation vector pCGN1559 (McBride and Summerfelt, 1990, *Plant Mol. Biol.* 14: 269-276) carrying a neomycin phosphotransferase gene under the control of the 35S CaMV promoter. The resulting plasmid was introduced in *Agrobacterium* (strain EHA101).

Transformation and Regeneration

Plants were transformed and regenerated as described in example 1. Transgenic plants were initially identified using a neomycin phosphotransferase assay and subsequently confirmed by expression of protein A fusions through immunoblot analysis.

Preparation of Oil Bodies

Oil bodies from the transgenic *B. napus* and *B. carinata* lines expressing the oleosin-protein A fusion were prepared following the procedure described in example 1.

Binding of Oleosin-Protein A Fusions to IgG

Oil body protein extracts (20 µg/ aliquot) from various transgenic *B. napus* lines expressing oleosin-protein A fusion proteins were subjected to polyacrylamide gelelectrophoresis and subsequently transferred to a PVDF membrane following standard procedures. The membrane was then probed with a HRP-conjugated mouse anti-rabbit antibody and visualised following the procedure as outlined in Antibodies, a laboratory manual (Harlow and Lane, 1988, Cold Spring Harbor). In Figure 16 the stained PVDF membrane is shown. A 50 kDa protein (predicted molecular mass of the oleosin-protein A fusion protein: 48,801 Da) was specifically detected in the protein extracts of all of the six transgenic *B. napus* lines tested. Untransformed control plants did not exhibit HRP activity, while the a 30 kDa protein (predicted molecular mass 29,652 Da) was present in a bacterial lysate transformed with pRIT2T encoding protein A and undetectable in the untransformed lysate.

Binding and Elution of IgGs t Oil Bodies Containing Oleosin-Protein A Fusi n Proteins

Washed oil bodies (10 mg/ml protein) were prepared from wildtype *B. napus* and a transgenic *B. napus* line transformed with a construct expressing an oleosin-protein A fusion protein as described in example 1 and suspended in 10 mM Tris-Cl pH 8.0. A volume of 2 µl (±34 µg) of HRP-conjugated rabbit anti-mouse antibodies (Sigma, cat no A9044) was added to 500 µl of the washed oil body preparation and the suspension was incubated for 1 hr at room temperature or overnight at 4°C. The samples were then centrifuged for 15 min at 16,000 xg and the undernatant was removed. Subsequently, the oil bodies were thoroughly resuspended in 500 µl 10 mM Tris-Cl pH 8.0 using a pestle. This washing step in Tris-Cl was repeated 4 times (henceforth termed washed oil body preparation). A 5 µl aliquot from the washed oil body preparation was washed a fifth time and then assayed for HRP activity.

HRP assays were carried out by adding 1 µl of the washed oil body preparation to 1 ml of HRP assay mix (9.8 ml of 0.1 M NaOAc, 0.2 ml of 2.5 mg/ml Trimethylbenzidine in DMSO, 4 µl H_2O_2) and incubating the mixture for 5 min at room temperature. The reaction was then stopped by adding 0.5 ml 1M H_2SO_4. The samples were filtered through a 0.22 µm filter and subsequently the OD_{450}'s were determined spectrophotometrically.

In order to elute the IgGs from the oil bodies, the washed oil body preparation was resuspended in 100 mM glycine pH 3.0 and centrifuged for 15 min at 16,000 xg and incubated for 30' at room temperature. Following neutralization in 500 µl 100 mM Tris-Cl pH 8.0, both the oil body fraction and the eluate were assayed for HRP activity as above. The binding and elution of IgGs to oil bodies from weld type *B. napus* and transgenic *B. napus* expressing an oleosin protein A fusion, are illustrated in Figure 17.

All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

EXAMPLE 10

Construction of a PRS-OBScFv-prochymosin plant expression vector

The example below describes the construction of a plant expression vector containing a gene sequence which upon expression in the plant produces a fusion protein comprising a signal sequence, a Single-chain variable Fragment (ScFv) isolated from a mouse hybridoma cell line producing monoclonal antibodies raised against an *Arabidopsis* oleosin and the zymogen chymosin.

Arabidopsis cDNA clone Atts0278 containing an *Arabidopsis* oleosin cDNA sequence was obtained from the *Arabidopsis* Biological Resource Centre (ABRC, http://aims.cps.msu.edu) *Arabidopsis* stock centre. This cDNA sequence is identical to the coding sequence of the *Arabidopsis* oleosin genomic sequence as published in (Van Rooijen et al (1992) Plant Mol. Biol. 18: 1177-1179). Using standard molecular biology techniques such as polymerase chain reaction (PCR), ligation and transformation, this sequence was furnished with restriction sites which allowed for the in-frame cloning of this sequence in the bacterial expression vector pRSETB (obtained from Invitrogen). Manufacture's (Invitrogen's) protocols were used to express the His-tagged *Arabidopsis* oleosin. The His-tagged *Arabidopsis* oleosin protein was run on a polyacrylamide gel, electroeluted and used to obtain mouse hybridoma monoclonal antibodies using standard laboratory techniques (See eg Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989). A mouse hybridoma cell line which produces a monoclonal antibody which specifically recognizes the *Arabidopsis* oleosin described above (and very similar Brassica oleosins) was called D9.

The hybridoma cell-line was used as a source of mRNA for the production of a single-chain variable fragment (ScFv) gene in which the variable regions of the antibody heavy (Vh) and light (Vl) chain genes are joined by a flexible peptide linker. This mRNA was isolated using standard laboratory techniques. The production of the anti-oleosin ScFv gene was achieved utilizing a "Recombinant Phage Antibody System Mouse ScFv Module" obtained from Pharmacia Biotech (Code No: 27-9400-01). The OBScFv gene will be referred to as the Oil body Single-chain variable Fragment (OBScFv) gene. The OBScFv gene was furnished with an ER signal

sequence termed PRS at the 5' end of the OBScFv gene to allow for secretion into the plant cell apoplast upon expression of this gene in plants. This was accomplished using gene splicing by an overlap extension technique (Horton et al GENE (1989) 15: 61--68. The PRS DNA sequence was chemically synthesized and encodes a signal sequence which is identical to the deduced amino acid sequence of a tobacco E2 thaumatin-like protein [Van Kan et al (1989) Plant Mol Biol 12: 153-155]. The 3' end of the OBScFv gene was fused in frame to a sequence encoding the bovine zymogen prochymosin using gene splicing by overlap extension technique (Horton et al GENE (1989) 15: 61--68). The gene sequence encoding this prochymosin was codon optimized for expression in plants. The deduced aminoacid sequence of the prochymosin is identical to the deduced aminoacid sequence of prochymosin cDNA as reported in Harris et al (1982) NAR 10: 2177-2187.

The PRS-OBScFv-Prochymosin gene fusion as described above was placed under the regulatory control of the phaseolin promoter and the phaseolin terminator derived from the common bean *Phaseolus vulgaris* (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-Gopalan *et al*., (1985) PNAS USA 82: 3320-3324)). A gene splicing by overlap extension technique (Horton et al GENE (1989) 15: 61--68) was used to fuse the phaseolin promoter to the PRS-OBScFv-Prochymosin gene. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were used to furnish the phaseolin promoter and terminator with Pst I and HindIII/KpnI sites respectively (see Figure1). Standard molecular biology laboratory techniques were also used to place the phaseolin terminator dowstream from the PRS-OBScFv-Prochymosin gene. The PstI-phaseolin promoter-PRS-OBScFv-Prochymosin-phaseolin terminator-KpnI insert sequence was cloned into the PstI KpnI sites of pSBS3000 (pSBS3000 is a derivative from the *Agrobacterium* binary plasmid pPZP221 (Hajdukiewicz et al., 1994, Plant Mol. Biol. 25: 989-994). In pSBS3000, the CaMV35S promoter-gentamycin resistance gene-CAMV 35S terminator of pPZP221 was replaced with parsley ubiquitin promoter-phosphinothricin acetyl transferase g ne-parsley ubiquitin termination sequence to confer resistance to the herbicide glufosinate ammonium. The resulting plasmid is called pSBS2168. The

sequence of the phaseolin promoter- PRS-OBScFv-Prochymosin-phaseolin terminator sequence is shown in Figure 1.

Plasmid pSBS168, was electroporated into *Agrobacterium* strain EHA101. This *Agrobacterium* strain was used to transform *Arabidopsis*. *Arabidopsis* transformation was performed essentially as described in "*Arabidopsis* Protocols; Methods in molecular biology Vol 82. Edited by Martinez-Zapater JM and Salinas J. ISBN 0-89603-391-0 pg 259-266 (1998) except that the putative transgenic plants were selected on agarose plates containing 80μM L-phosphinothricine, and subsequently transplanted to soil and allowed to set seed.

EXAMPLE 11

Polyacrylamide gelelectrophoresis (PAGE) and immunoblotting of transgenic seed extracts.

Preparation of total Arabidopsis seed extracts for PAGE. 50 mg of transgenic *Arabidopsis* seeds were ground in 1 ml of oil body extraction buffer containing 0.4 M sucrose, 50 mM Tris-Cl pH 7.5 and 0.5 M NaCl. A 25 μl aliquot was taken out and labeled "total extract". The rest of the extract was placed at room temperature for 20-30 minutes to allow the PRS-OBScFv-Prochymosin protein to associate with the *Arabidopsis* oil bodies. This mixture was spun at room temperature for 10 minutes at 10,000 x g to allow for the separation of the oil body fraction (floating on top) from the supernatant (25 μl was taken out and labeled Sup0,) and the pellet fraction. The oil body fraction was resuspended in 300 μl oil body washing buffer containing containing 0.4 M sucrose, 50 mM Tris-Cl pH 7.5. This fraction is called Obtotal.

25 μl of OBtotal was mixed with 25 μl of a high stringency washing buffer (8M urea, 100mM Na2CO3) spun for 10 minutes at 16,000 x g . The oil body fraction was resuspended in 25 μl high stringency washing buffer ... 5 minutes and the oil body fraction was resuspended in 25 μ

The "total extract", Sup 0, OB0, and OBHighS samples were prepared for SDS PAGE and Western blotting using standard molecular biology protocols. Equal amounts of sample were loaded in each lane of the gel (As the total extract has a volume of 1ml and the oil bodies are resuspended in approximately 300µl, the oil body associated proteins are enriched about 3-fold in lane 4 and 5 compared to lane 2). The results of this experiment are shown in figure 2. As can be seen from this figure the majority of the OBScFvProchymosin protein fusion has been autocatalytically processed into mature chymosin and several proteolytic breakdown products of chymosin (here referred to as "partial chymosin". This autocatalytical processing is typical for secreted chymosin fusion proteins (personal observations and Ward et al (1990) Bio/Technology 8: 435-440). As seen in lane 2, in addition to mature and partial chymosin, two bands of the predicted molecular weight of an intact OBScFvProchymosin fusion protein can be found. The top band has a slightly lower mobility on a polyacrylamide gel and does not separate with *Arabidopsis* oil bodies (lane 4 and 5). Instead this band is found in the supernatant fraction (lane 3). It is predicted that this is a suboptimally folded form of the OBScFvProchymosin fusion protein. The mature and partial chymosin products are not expected to co-purify with the oil bodies as they have been cleaved from the OBScFv "carrier". As can be seen in figure 2 the ratio of OBScFvProchymosin to mature chymosin is dramatically increased in the oil body samples (lane 4 and 5) which indicates that the OBScFvProchymosin protein has a higher and specific affinity for oil bodies. When a high stringency urea buffer is used (lane 5), the only protein which still co-separates with the oil bodies is the OBScFvProchymosin fusion protein. This confirms that this binding is specific and that the OBScFv portion of the OBScFvProchymosin protein is responsible for this binding. Equal amounts of sample were loaded in each lane of the gel.

INS B2